# CHAPTER 8

# TESTING

## INTRODUCTION

Navy schools use tests to determine whether or not a student has sufficient knowledge or skill to meet the requirements established by the learning objectives; that is, whether or not the student has learned the material. The philosophy underlying Navy testing is based on the achievement of learning objectives. Tests are given to determine if a student can demonstrate, in some measurable way, achievement of the objectives.

You will fill a critical role in the testing program for the courses you instruct. After curriculum has been validated, course personnel (primarily instructors) are responsible for the development of additional versions of tests, development of additional test items, and analysis of tests and test items.

You will be concerned with two methods of testing; knowledge and performance. Knowledge tests measure achievement of objectives through the use of test items written at the appropriate learning level. Performance tests measure skill acquisition by having the student demonstrate specific behaviors defined by the learning objectives. This chapter focuses primarily on the information you will need to develop knowledge test items.

## KNOWLEDGE TEST ITEM DEVELOPMENT

The behavior, conditions, and standards specified in the objectives will determine the level of learning tested. You need to know how students will use this material in the job so that you can test the material to that level. Navy training uses five levels of learning which are based on, though not identical to, the learning levels defined in Chapter 7. Definitions and examples of the five learning levels are as follows:

**Recognition.** Recognition is the process of verbatim identification of specific terms, facts, rules, methods, principles, procedures, objects, and the like, presented during training. Students select from two or more alternatives to identify the information. For example, a test item may ask the students to identify a particular switch on a piece of equipment by matching its name to a diagram of the switch. That is a recognition test item if the student has been taught that specific information during training.

**Recall.** Recall is the verbatim remembering of specific terms, facts, rules, methods, procedures, principles, and the like. To correctly answer a recall test item, students remember and respond exactly as taught. A recall test item requires students to respond from memory instead of selecting the response from two or more alternatives. Listing the steps of a maintenance procedure and answering a completion question by labeling parts on a diagram are

examples of recall test items. Always test recall with closed book tests, otherwise you are not testing the students' ability to remember information.

**Comprehension.** Comprehension is understanding what was taught rather than simply memorizing the words. It can be demonstrated by interpreting, explaining, translating, or summarizing information. When measuring the students' understanding of an objective, you must avoid the use of verbatim recall or recognition types of items. Comprehension requires you to paraphrase the material presented in the item rather than taking it word for word from the text. Asking a student to explain how a device works is an example of a comprehension item.

**Application.** Application involves the ability to use acquired knowledge in a job-related situation. Application questions require students to demonstrate knowledge through mental skill exercises such as solving a computational problem or determining resistance values from circuit diagrams. You must use different problems or circuits from the ones you used in class to develop application questions.

**Analysis/Evaluation.** Analysis involves the understanding of the elements of data and relationships among the data that make the meaning of information explicit. Evaluation involves the judgment of the value or the effectiveness of procedures or solutions based on data, criteria, and standards. For example, consider a question that asks the student to select the best approach to meet a stated objective. The question would require the student to know or determine which options would meet the objective (analysis) and which single option would be best (evaluation).

In developing knowledge test items, focus on the learning level being tested and write the test items to that level. You may use five types of knowledge test items: multiple-choice, true-false, matching, completion, and essay.

## MULTIPLE-CHOICE TEST ITEM DEVELOPMENT

The multiple-choice item is the most versatile of the five types of test items. Use it to test all levels of knowledge except recall. The multiple-choice test item consists of (1) a stem containing the problem statement and (2) a list of possible answers (alternatives).

Typically, this type of test item contains four alternatives; however, depending on the nature of the content being tested, you can use more or less than four. Make one of the alternatives the correct answer to the test item and all of the others plausible alternatives.

The following sections present guidelines for stem construction, alternative construction, test item forms and formats, and common errors in item construction.

### Stem Construction

A cardinal rule in test item development is to communicate effectively. Use the following guidelines as a checklist to make sure you properly write multiple-choice test item stems:

- Include all information, conditions, assumptions, and details required for the students to correctly answer the question without requiring them to refer to the alternatives.

- Phrase the stem positively.
- If you must use a negative, highlight it (in caps or underlined) so that the student will notice it and interpret the item correctly.
- Use clear, unambiguous wording so that only one answer is correct.
- Include words, phrases, and so on, that pertain to all alternatives rather than repeating them in the alternative.
- Omit information not essential to the interpretation of a test item.
- If the test item uses an illustration on a separate illustration form, refer to the illustration in the stem by figure number.
- Use complete sentences for test items in the form of questions and end them with a question mark.
- Position the completion of an incomplete statement test item near or at the end of the stem.
- Avoid the use of more than one completion position.
- Use the question form over the incomplete statement form except when it would make the test item grammatically clumsy or difficult to understand.
- Test only one idea or central thought.

**Alternative Construction**

You must exercise care when designing the alternatives for multiple-choice test items. Make sure the alternatives are plausible and fit well with the stem. The difficulty of the item will depend largely upon the alternatives. The more closely related the alternatives are, the more difficult it is for students to select the correct answer. A good rule is to develop alternatives based upon common misconceptions by students and inexperienced job incumbents. You may prepare alternatives based on how students might incorrectly manipulate terms, symbols, and the like. An additional rule is to look at the correct answer and determine how you may make it incorrect. Observe the following requirements in developing multiple-choice alternatives:

- Include only one correct answer.
- Use closely related alternatives.
- Use alternatives that are meaningful and not subject to automatic elimination because they are irrelevant or unrelated to the question.
- Do not use interrelated answers (e.g., “c” is true if “a” and “b” are false).
- Use terms with which students are familiar or that you can explain within the limits of the test item.
- Make all alternatives approximately the same length and of the same complexity.
- Do not use the words “always” or “never.”
- Do not use alternatives of “all of the above” and “none of the above.

- Avoid using negative wording. However, if you must use negative wording, highlight it (e.g., in caps or underlined).
- Punctuate alternatives to conform grammatically with the structure of the item stem.
- If the stem is a question (i.e., a closed-stem) and the alternatives are complete sentences, begin each alternative with a capital letter and end each one with a period.
- If the stem is a question and the alternatives are incomplete sentences, begin each alternative with a capital letter and use no end punctuation.
- If the stem is an incomplete sentence (open-stem) with the response position at the end of the stem, begin each alternative with a lowercase letter (except for proper nouns) and end it with a period.
- With the incomplete sentence test item, make the wording of the alternatives grammatically related to that of the item stem.
- Randomly select the position of the correct answer among the alternatives to avoid any patterns that may bias the test.
- In items that involve numerical answers, arrange the alternatives in ascending or descending order.

**Multiple-Choice Stem Formats**

You will use two formats to construct the stem of multiple-choice test items; the *closed* and the *open*.

Closed stem format. You may write closed stem items as a complete statement or incomplete statement. The following is an example of a complete statement format:

EXAMPLE: Which of the following actions is required to remove a hinged type 2 module on the MTRE Mk 7 Mod 2/4?

(a) Disconnect plates from the type 2 module.

(b) Insert “T” handle into quick release fasteners.

(c) Remove all Type 3 modules and connectors.

(d) Rotate hold down clamps to a vertical position.

The complete statement format has the advantage of forcing you to state the problem clearly in the stem. It also reduces the possibility of giving students grammatical clues. A disadvantage is that it may require lengthier responses. The following is an example of an incomplete statement format:

EXAMPLE: The setting of the AN/ABC-3Q flip-flop . . . indicates that intent-to-fire has been energized.

(a) B43

(b) C21

(C) C24

(d) D32

When written as an incomplete statement, the completion position appears within the statement, not at the end of the stem. Although this form of test item is typically easier to write than complete statement stems, use them sparingly. They encourage lifting of test items verbatim from the material and encourage students to memorize answers.

Open stem format. This format uses an open-ended stem, which is an incomplete statement with the response position at the end of the statement. Each alternative provides a logical conclusion to the stem. Although incomplete-statement stems are typically easier to write than complete statement stems, they may cause you to avoid thinking about the question before you develop the alternatives. That may result in illogical and unrelated alternatives. Generally, the less similar alternatives are in content, the easier it is for students to select the correct alternative. The following is an example of an open stem test item.

EXAMPLE: When crimping both a stranded and a solid wire in the same contact, the solid wire's position in relation to the stranded wire is

(a) above.

(b) below.

(c) beside.

(d) diagonal.

**Multiple-Choice Test Item Formats**

You may construct a multiple-choice test item either as a question or an incomplete statement using the *standard* or *except* formats.

Standard Format. This particular format is straightforward and the easiest to develop. Use it when you only want students to select the correct answer from the four alternatives provided.

EXAMPLE: During the system verification test, what supplies voltages for TVC position sensor tracking?

(a) Minus 20 VDC precision power supply.

(b) Self-test DC reference power supply.

(c) TVC position sensor AC/DC converter.

(d) Missile command module.

Except Format. Use the except format when three or more equally correct alternatives answer the question. This format requires students to recognize which alternatives are correct and to select the one that is incorrect. Always capitalize or underline the word "EXCEPT" in the stem. Use the "EXCEPT" format sparingly.

EXAMPLE: A specific torquing pattern and associated torque values can be found in the SINS technical manual for all of the following assemblies or components EXCEPT

(a) An azimuth synchro assembly mounted to the stem.

(b) A velocity meter mounted to the platform.

(c) A replacement gyroscope mounted to the stable platform.

(d) A platform stem mounted to the bedplate.

There are several common errors that you need to avoid when developing multiple-choice test items. Listed below are four examples of common errors:

Do NOT use similar wording in both the stem and ONLY the correct alternative. It suggests the correct answer.

Example of an inappropriate test item: (error underlined):

What is the purpose of the MARDAN maintenance test set?

(a) Monitors the C.P. operations.

(b) Furnishes power to MARDAN.

(c) Functions as a running time meter.

(d) Provides static testing of MARDAN.

Do NOT state the correct alternative in greater detail than the other alternatives. This practice often cues the correct answer.

Example of an inappropriate item (error underlined):

When all weapon power is removed from the PIP, which of the following statements is true?

(a) All power is lost to the MCC equipment.

(b) The MCC equipment is furnished power from NAV via the MSR.

(c) The DCCs have heater power applied.

(d) Power from the ship control center may be present in MCC since it only goes through the SHIP JP.

Do NOT use two or more alternatives that have the same meaning. It eliminates them as useful alternatives and simplifies the choice. In the following example, alternatives 1 and 2 have the same meaning. Thus, they reduce the number of realistic alternatives from three to one.

Example of an inappropriate item (error underlined):

What is the final step in performing post maintenance checks?

(a) Secure the front panel to the chassis.

(b) Make sure the front panel is secure.

(c) Set manual test switch to "OFF."

(d) Rerun the diagnostic tests.

Do NOT use alternatives that are included in other alternatives. In the following example, alternative 2 includes alternative 1. If alternative 2 is correct, then so is alternative 1.

Example of an inappropriate item (error underlined):

What is the operating time, in seconds, for the pressurization/compensation blow valve to roll from shut to open?

(a) 1 to 3

(b) 1 to 4

(C) 4 to 6

(d) 9 to 11

## TRUE-FALSE TEST ITEM DEVELOPMENT

The true-false item is a two-response multiple-choice item. Use it only when one plausible alternative to an item exists. A major drawback to the true-false item is that it is susceptible to guessing. A student who does not know the correct answer has a 50-percent chance of responding correctly to the item. Use true-false items to test recognition, comprehension,

application, or evaluation. Use the following guidelines when writing the true-false test item.

### True-False Test Item Format

The true-false item format is straightforward. Write the stem as a direct statement and label the two alternatives "true" or "false."

EXAMPLE: (TRUE/FALSE) When placing the CA in stowage, you must make sure the CA temperature is normal before securing heater power.

a. True

b. False

### True-False Test Item Construction

Observe the following rules for constructing true-false items:

- Include all relevant information and conditions required for the students to correctly answer the item in the descriptive statement.
- Make the statement concise and clear. Make sure the proposition that makes the statement true or false is evident.
- Make sure the statement is clearly true or false.
- Place the TRUE/FALSE identification before the item.
- When possible, make a false statement consistent with a typical misconception.
- Do not use specific determiners (e.g., always, never, none, all, may, sometimes).
- Keep items short. Long items are harder to read and more difficult to judge true or false.
- When possible, use positive statements to minimize confusion.
- Do NOT lift test items verbatim from the curriculum.

## MATCHING TEST ITEM DEVELOPMENT

The standard matching format consists of two lists containing related words, phrases, or symbols. Students must match elements on one list with associated elements on the other list based on specific instructions. Students pair the elements in each list and record the answer. Matching test items are ideally suited for testing recognition but may also test comprehension and application.

### Test Item Format

The matching test item consists of a stem and two columns listed below the stem. The stem provides directions on how the student must match the items in the two columns. One column contains the questions or problems to be answered and the other column consists of the answers.

EXAMPLE: Using the FCDs in OP XXXX, match the circuit element listed in column B to the signal it generates in (column A). Write the letter representing your answer in the blank to the left of each signal in column A. You may use a letter in column B once, more than once, or not at all.

| COLUMN A | COLUMN B |
|---|---|
| 1.___DATA CHK NOT OK | a. B10 |
| 2.___DATA CHK OK | b. B13 |
| 3.___DRY RUN | c. B16 |
| 4.___EQ CONT RST 2 | d. B46 |
| 5.___DATA CHK REQ | e. B49 |
| 6.___DATA CHK ALM | f. C30 |
| | g. D56 |

**Test Item Construction**

Use the following guidelines when constructing matching test items:

- Clearly specify in the stem (directions) how the students are to match the question and the answer.
- Always place the questions in the left-hand column. Place answers in the right-hand column.
- When feasible, use single words, numbers, codes, symbols, short phrases, and the like, in the answer list.
- Make all answers relate to the question. That helps to prevent elimination of unrelated answers.
- Specify in the directions how often students may use the answers.
- When possible, arrange the answers according to some system (e.g., arrange numerical answers in ascending or descending order).
- Place options on the same page. Students should not have to turn back and forth for the answer.

## COMPLETION TEST ITEM DEVELOPMENT

The completion test item is a free response type of item in which the student must supply the missing information from memory. You may make the completion item a listing test item in which the student must supply the required list of part names, procedural steps, and so on, from memory. An advantage of the completion item over the multiple-choice or the true-false types is that it requires more than simple recognition of information. That eliminates the possibility of guessing.

Completion items are easy to construct. You will find them useful in situations in which

students must write a computational equation, define terms, list part names and functions, and the like. The disadvantage is that it is more difficult to score and must be accompanied by grading criteria.

### Test Item Format

You can construct completion items using three basic formats:

- Students supply the word or phrase that completes the statement.

  EXAMPLE: The station clock and time display tests check the performance of the individual stages of the register designated . . .

- The student provides a definition, term, formula or similar response to a question.

  EXAMPLE: What is the name of the unit that detects angular motion and supplies an output through precession?

- The student supplies a list of procedures, steps, and so forth, from memory. This type of test item may be expressed in question or statement form.

  EXAMPLE OF STATEMENT FORM: In the space below, list in order the steps for placing the GA in stowage.

  EXAMPLE OF A QUESTION FORM: What are the steps for performing a MARDAN maintenance test?

### Test Item Construction

Use the following guidelines when constructing completion items:

- Word the test item clearly and comprehensively enough to allow a student who is knowledgeable in the subject area to answer correctly.
- Make sure the missing segment of the incomplete statement item is important, such as a key element of a process or a piece of equipment.
- In incomplete statement items, do not omit too many words or the statement will become unclear and force students to guess.
- In incomplete statements, make sure the response position appears near or at the end of the stem. Items with the response position near the beginning are harder to read and take longer to answer.
- Provide sufficient space on the answer sheet for students to enter their response.
- Use a direct question to test for comprehension of technical terms or knowledge of definitions.

- Do not make the correct answer a "giveaway" word that could be guessed by students who do not really know the information. In addition, avoid giving grammatical cues or other cues to the correct answer.
- Avoid using statements taken directly from the curriculum.
- Develop grading criteria that lists all acceptable answers to the test item. Have subject matter experts determine the acceptable answers.

## ESSAY TEST ITEM DEVELOPMENT

The essay test item requires the student to answer a question with an original written response. Use comprehension essay test items for testing the student's ability to organize data and express thoughts clearly in writing. Do not use them to test recall. Essay tests involve a relatively subjective scoring process since many factors may enter into the correctness of a response.

The disadvantage to essay test items is that they are time-consuming and difficult to score. The essay item must be scored by an individual knowledgeable in the subject area, unless only one basic response is possible to a given question or requirement,

### Test Item Format

You can use an essay question to assess learning of a comparatively large body of information, as well as individual elements within that body. Use the following guidelines for formatting the essay test item:

- State clearly and precisely what type of response is required.
- If possible, place limits on the response by identifying the major points the students should address, the length of the response required, or time students may spend on the response.

EXAMPLE: Compare the gas turbine and the 1200 PSI propulsion plants. Your discussion should include descriptions of the major components of each system. Partial credit will be given.

### Test Item Construction

The following are examples of types of information for which you might want to use essay test items:

- A comparison or contrast of items and procedures.
- A decision for or against system or equipment operation.
- Relationships such as causes and effects.
- Illustration (sketch) of principles learned.
- Statement of purpose in the selection of a method or technique.
- Criticism of the adequacy or correctness of a diagram or procedure.

- Discussion of primary, alternate, or emergency procedures.
- Explanation or definition of tasks.
- Observation from illustration or operation.
- Evaluation of the appropriateness of a procedure or technique.

### Model Answer or Grading Criteria

The essay test item must also contain a model answer you will use to grade the question. Observe the following guidelines in developing a model answer:

- Make sure the grading criteria identifies all of the essential information a knowledgeable student should be able to supply.
- Make sure it promotes objective scoring of the test item by establishing a standard answer from which to judge all others.
- Make sure it identifies how much each item or part of an item is worth.

### Validation Of Test Items

Once you have constructed the test items, and before you actually assemble the test, validate the content of the items. Make sure they are technically and grammatically accurate, that they measure the objective, and that the items adhere to the guidelines presented in the preceding paragraphs. Have technically qualified SMEs perform the validation process. The individuals validating the test item should answer the following questions:

- Is the item technically accurate and is the correct response keyed?
- Is the item written to measure the objective?
- Does the item measure a knowledge critical to the task associated with the objective?
- Is the item written to the appropriate learning level?
- If recognition, recall, or comprehension of the knowledge being tested is required for competent performance on the job, is the item a closed-book item?
- If the knowledge being tested is normally looked up during performance of on-the-job task(s), is the item an open-book item and is the essential reference material supplied?
- Are all words spelled correctly? Is the grammar correct?
- Does the item meet format construction guidelines?

If the answer to any of the preceding validation criteria is NO, correct the discrepancy and revalidate the test item. If the item meets the validation criteria, then it should be approved for use.

### Test Item Analysis

After the test items have been reviewed for content validity and administered to the students, statistics will be kept by the course personnel to complete the validation process. These statistics include discrimination; difficulty; and for multiple choice items, effectiveness of alternatives.

Curriculum development manuals NAVEDTRA 130 and 131, in addition to NAVEDTRA 135, contain additional information on tests and test-item analysis.

## PERFORMANCE TEST DEVELOPMENT

The goal of many courses, as reflected in the learning objectives, is to teach students to perform skills needed on their job. Therefore, performance testing will constitute a significant portion of the testing for many courses. Objectives that require the demonstration of observable skills are tested by performance tests. Performance tests include the following considerations:

- Performance tests are simulated work situations in which students demonstrate their ability to complete procedures, produce a product, or a combination of both.
- Evaluation of performance usually involves the detailed observation and critique of a student's performance by a trained evaluator or instructor. The evaluation is supported by checklists or rating scales.
- The performance is observed and evaluated under the conditions and standards set forth in the learning objectives.
- A final product performance test involves comparing the student's efforts to an acceptable completed example.
- Skill objectives to be performance tested are identified and rated as to their relative importance in measuring student attainment of the related job skills.

Development of performance tests can be very involved. Performance tests consist of a simulated work situation in which the student performs a task based on a skill objective. Two types of performance tests are used to measure skill achievement: process and product. The development steps are essentially the same for both types with the exception of the final evaluation device. Some performance tests require a combination of both process and product measurement. NAVEDTRA 130 and NAVEDTRA 131 provide detailed information about the development of performance tests and their elements.

## SUMMARY

Assessing a student's mastery of objectives in a knowledge test or skill acquisition in a performance test is a natural progression in the learning process. For the assessment to be accurate, test items should be directly related to the level of learning you want to measure and they should be validated and analyzed by subject matter experts. Instructor Training School uses both types of tests, and students are able to experience firsthand the benefits and limitations of each.